The EMBO Journal vol.6 no.1 pp.133–138, 1987

Expression of the human papillomavirus type 18 E7 gene by a cassette-vector system for the transcription and translation of open reading frames in eukaryotic cells

Hans-Ulrich Bernard, Tilman Oltersdorf[1] and Klaus Seedorf

Institute of Cell and Tumor Biology and [1]Institute of Virus Research, German Cancer Research Center, Im Neuenheimer Feld 280, D-6900 Heidelberg, FRG

Communicated by G.Schütz

We have constructed and functionally tested a cassette-vector-system for the transcription and translation of open reading frames (ORFs) in cells of higher eukaryotes. The vectors are derived from the plasmid pBR322 and can be selected and amplified in *Escherichia coli*. Alternative eukaryotic promoters can be inserted between the restriction sites *Sph*I and *Kpn*I, translation initiation motifs between *Kpn*I and *Bgl*II, linkers for the adjustment of the translation reading frame and the insertion of genes or gene segments between *Bgl*II and *Hin*dIII, followed by a *Hin*dIII – *Eco*RI segment with splicing and polyadenylation signals derived from SV40. A prototype vector system, pORFEX11, 12 and 13, contains the strong cytomegalovirus immediately early promoter and a 10-bp motif of the SV40 T-antigen translation start. Polylinkers derived from pUC18 permit the insertion of ATG-less ORFs downstream from the ATG of the vector. Either of the three alternative polylinkers adjusts the appropriate translation frame. A similar construct contains the regulatable promoter of the *Drosophila* heat shock gene 70. We inserted genes or gene segments, that code for the bacterial chloramphenicol acetyltransferase, the bacterial gene conferring resistance against hygromycin, and the ORF E7 of the human papillomavirus type 18 into these vectors. After transfection of mouse L fibroblasts, all proteins and functions were expressed in accordance with the prediction. In transiently transfected L cells, the E7 protein expressed from pORFEX12 constitutes ~2.0% of total cell protein. This E7 protein could be localized by immunocytochemistry as a cytoplasmic component.

Key words: human papillomavirus/expression vector/transfection/translation initiation

Introduction

It has become a standard technique for genetic engineering to fuse a well-defined promoter sequence with a DNA segment that codes for a gene product of interest. After transient or stable transfection of cell cultures the gene products, mRNA and protein, are expressed and can be analyzed or their function exploited. This approach permits the expression of bacterial genes in eukaryotic cells (Mulligan and Berg, 1980), the separation of genetic units of complex viral genomes (Nakabayashi *et al.*, 1983), the expression of genes whose own promoter is limited to function in cells of a particular differentiation type (Hsiung *et al.*, 1984), or the expression of genes on cDNA clones (Okayama and Berg, 1982). This technique finds its limits when open translation reading frames (ORFs) on cDNA clones, or on viral or genomic exons, lack their genuine translation initiation signals, most of all their ATG codon. These gene sequences can only be expressed by fusing the available fragment to the N-terminal part of a well characterized protein. To plan and construct such a fusion gene is normally very time consuming and leaves the experimenter with fusion genes and fusion proteins of poorly predictable characteristics.

To overcome these technical limitations we have constructed a vector system and tested the predictions about its properties. We found it particularly desirable to have a versatile system where singular restriction sites flank the relevant elements of the vector and make their separate exchange feasible (Figure 1). Thus, any promoter can be inserted in the form of an *Sph*I – *Kpn*I fragment. This can be easily obtained by excising it from pUC18 (Yanisch-Perron *et al.*, 1985) after initially inserting it into the *Sma*I site of this plasmid. An adjacent 73-bp segment contains an ATG within a motif designed according to the translation initiation motif of the SV40 T-antigen gene and a polylinker, which was modified by insertion of one or two additional bases to give all three translation fragments relative to the ATG. Nine singular restriction cuts in this segment permit a new design of the translation initiation motif, modifications of the N-terminus of the expressed protein and the insertion of ATG-free or ATG-containing genes and gene segments.

We approached this technical objective with an interest in the functions of the proteins coded by the human papillomaviruses (HPV) type 6, 11, 16 and 18, which are associated with benign and malignant neoplastic lesions most often found in the genital tract of man (for a review see Gissmann, 1984). In unpublished experiments we found that after transfection of these viral DNAs in various mouse or human cell lines no or only aberrant trans-

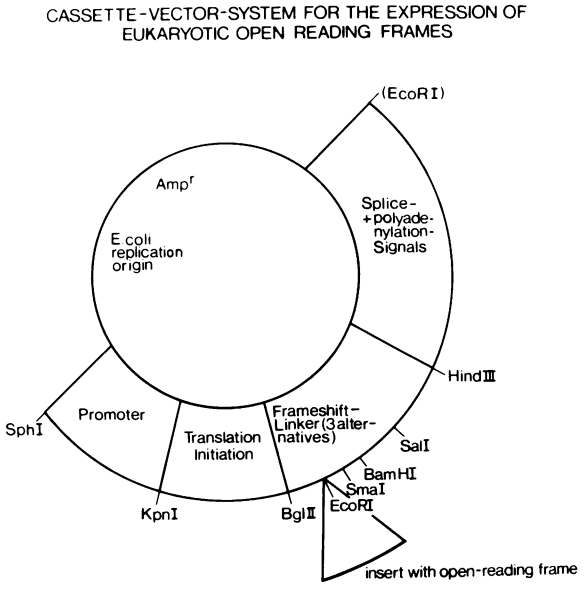

Fig. 1. Conceptual design of a cassette-vector system.

cripts can be obtained (H.U.Bernard, unpublished) and no proteins are detectable with antibodies raised against HPV-fusion proteins made in *Escherichia coli* (Seedorf *et al.*, 1987). Our vector system was designed to make each HPV-protein individually expressable. We show data of overexpression of E7, one of ~10 ORFs of papillomaviruses, which opens an avenue to study the function of the gene products of these complex genomes (for a review see Pfister, 1984).

Results

Construction of pORFEX11, 12 and 13

p297 (P.Gruss, unpublished) is a 4212-bp plasmid (Figure 2) that consists of a 2295-bp segment derived from pBR322 coding for replication and selection properties in *E. coli*. A 323-bp segment with the early promoter elements of SV40 is linked through a *Bgl*II site to a 848-bp *Bgl*II – *Bam*HI segment with the SV40 t-splice and a 746-bp *Bam*HI – *Eco*RI segment with the late SV40 polyadenylation signal. These DNA elements are also present in commonly used plasmids of the pSV2-type (Mulligan and Berg, 1980). We eliminated the *Bam*HI site by fill-in with Klenow-polymerase and blunt-end ligation, and replaced the *Pvu*II – *Bgl*II segment with the SV40 promoter by a synthetic 34-bp oligonucleotide shown in Figure 2 to give a construct termed pATG. The oligonucleotide contains a 10-bp motif designed according to the SV40-T-antigen start. This is flanked 5′ of the ATG by a *Sph*I and a *Kpn*I and 3′ by a *Bgl*II and *Hin*dIII site. We isolated the immediate early promoter of the cytomegalovirus (CMV) (Boshart *et al.*, 1985) as a 656-bp *Hinc*II – *Ava*II fragment (with 55 bp downstream of the cap site) from plasmid pCM5028 (M.Boshart, unpublished) and the Drosophila heat-shock protein 70 promoter (hsp) (Steller, 1985) as a 335-bp *Eco*RI – *Hin*dIII fragment (encompassing 85 bp downstream of the mRNA cap site) from plasmid pUC8-hsp70 (B.Hovemann, unpublished). Both fragments were made blunt-ended with Klenow polymerase, inserted into the *Sma*I site of pUC19, and recloned from inserts with the appropriate promoter orientation as *Sph*I – *Kpn*I fragments into pATG to give the vectors pORFEX10 (CMV promoter) and pORFEX20 (hsp promoter). The orientation of the promoters had been determined with the help of three asymmetrically located *Bgl*I sites on the CMV-promoter fragment and an *Xho*I site on the hsp-promoter fragment (data not shown). Transcripts that initiate from these promoters do not contain an ATG translation initiation codon before reaching the ATG on the synthetic insert, a prerequisite for efficient use of this ATG in a eukaryotic translation system (Kozak, 1978). Any DNA segment that starts with an ORF can be inserted in the *Bgl*II and *Hin*dIII sites of pORFEX10 and 20 with the prediction that in cells transcription would occur from the vector promoter and translation from the vector ATG if the inserted gene comes in the appropriate translation frame.

Subsequently, we aimed to increase the versatility of these vectors by making fusions in all three translation frames possible and by increasing the numbers of useful restriction sites. To do this, we eliminated the *Sph*I – *Bam*HI segment of pORFEX10 and its singular *Eco*RI site by restriction enzyme cleavage, Klenow-polymerase fill-in and blunt-end ligation. Between the *Bgl*II and the *Hin*dIII sites of the resultant plasmids we inserted three alternative modified pUC18-polylinkers (W.Röwekamp, unpublished) whose sequences are given in Figures 2 and 6. Figure 6 depicts the possible use of translation frames and the consequences for the N-terminus of expressed proteins in the resultant vectors pORFEX11, 12 and 13 as exemplified for *Eco*RI or *Bam*HI insertions.

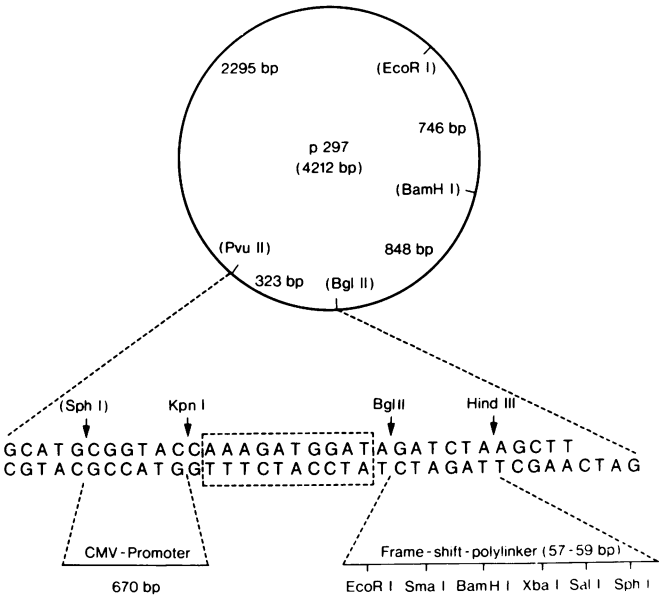

Fig. 2. Construction of pORFEX10 (no frame-shift polylinkers) and pORFEX11, 12 and 13. For a complete sequence of the frame-shift polylinker see Figure 6 and Yanisch-Perron *et al.* (1985). All restriction sites deleted in the final constructions are indicated by brackets.

Table I. CAT enzymatic activity in transiently transfected L cells

Plasmid	CAT converted (%)
–	0
pSV2CAT	12.4
pORFEX10-CAT	3.0
pORFEX10-CAT-ΔATG	60.9

Table II. Hygromycin-resistant clones after stable transfection of 10^6 mouse L cells with 1 μg of frameshift-ATG-Hm^r plasmids

Plasmid	Number of colonies
–	0
pHMR272	290
pORFEX10	0
pORFEX10-Hm16	0
pORFEX10-Hm19	11 (small)
pORFEX10-Hm24	220
pORFEX10-antisense-Hm24	0

Chloramphenicol acetyltransferase (CAT) enzymatic activity is strongly expressed by transcription from the CMV promoter

To obtain qualitative and quantitative information about the function of the CMV promoter on pORFEX10 and its derivatives, we decided to insert the gene coding for the bacterial CAT into this plasmid and to compare its expression with that from pSV2CAT (Gorman *et al.*, 1982) after transient transfection of mouse L cells with either of the plasmids. A *Hin*dIII – *Hpa*I fragment of pSV2CAT contains the complete CAT coding sequence plus 36 nucleotides upstream of its ATG and downstream of the gene a segment with the SV40 splice signal corresponding to the *Bgl*II – *Hpa*I segment of p297. This segment was inserted into pORFEX10, cut with *Hin*dIII and *Hpa*I, to give pORFEX10-CAT, a plasmid with an ATG on the vector in front of the CAT-

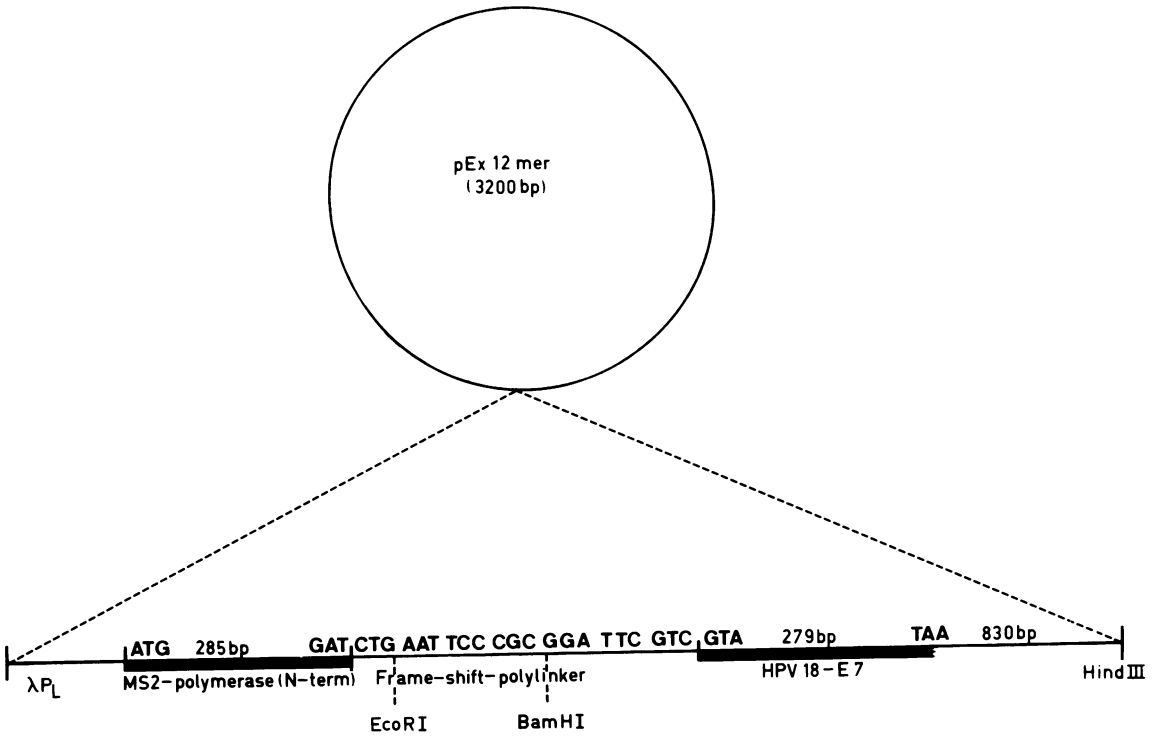

Fig. 3. Fusion of the HPV18-E7-ORF 3′ terminal to a sequence coding for the phage MS2-polymerase. This *E. coli* expression vector was used by Seedorf *et al.* (1987). In comparison to Figures 2 and 6, the sequence illustrates the in frame insertion of E7 as a *Bam*HI–*Hin*dIII fragment in pORFEX10, and as an *Eco*RI–*Hin*dIII fragment in pORFEX12.

ATG, but in a different translation frame. The ATG of pORFEX-10 was deleted from this construct by cutting pORFEX10-CAT with *Kpn*I and *Hin*dIII and by fusing the overhanging single strands with the oligonucleotide 5′-AGCTTGTAC-3′. This plasmid was termed pORFEX10-CAT-ΔATG. Table I shows that this plasmid stimulates about 5-fold higher CAT expression than the comparable construct based on the SV40 early promoter on pSV2CAT giving an estimate of the strength of the CMV promoter. The expression from pORFEX10-CAT, the plasmid with a second out-of-frame ATG in front of the CAT gene, is reduced by a factor of 20. The residual CAT activity most likely represents the infrequent use of the second ATG for translation initiation.

Gene expression from pORFEX10 is dependent on the translation frame

We have published the construction of *Bam*HI-linker insertion mutations in the region of the N-terminus of the bacterial gene coding for resistance against hygromycin (Hm) (Bernard *et al.*, 1985), work that led to the expression of this gene in mammalian cells under the influence of the Herpes simplex thymidine kinase promoter from the plasmid pHMR272. Three of these mutants contain the *Bam*HI linker in either of the three translation frames. We have inserted these three mutants as *Bam*HI–*Hin*dIII segments (Bernard *et al.*, 1985) into pORFEX10 cleaved with *Bgl*II and *Hin*dIII and transfected the resulting plasmids into L cells. Table II shows the colony count after two weeks of selection for hygromycin resistance. Only pORFEX10-Hm24 where the vector ATG and the Hm^r-gene are in the proper frame gives a number of resistant clones comparable to pHMR272. No clones were obtained with pORFEX10-Hm16, and only a few very small ones with pORFEX10-Hm19. The latter construct still contains the genuine ATG of the prokaryotic gene, but in a different frame from the ATG of the vector and the appearance of some small

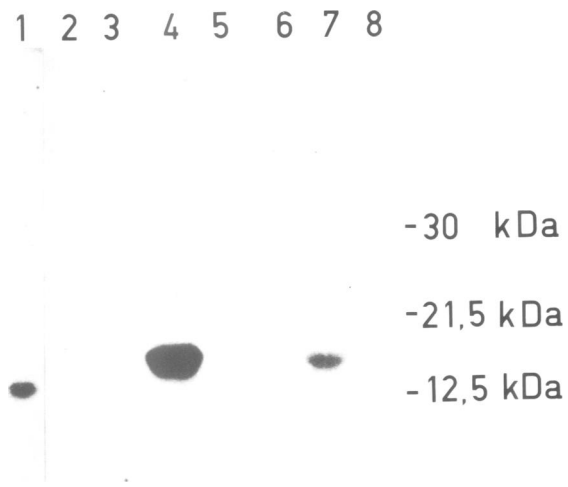

Fig. 4. Western blot analysis of protein from L cells (slots 2–7) transfected transiently with pORFEX-HPV18-E7 vectors, and of HeLa cell protein. The blots were probed with a rabbit anti MS2-HPV18-E7-IgG described by Seedorf *et al.* (1987). Autoradiography was done after reaction with ^{125}I-labeled protein A. **Slot 1**: HeLa cell protein; **slot 2**: no DNA; **slots 3–5**: HPV18-E7 in pORFEX11, 12 and 13; **slots 6** and **7**: pORFEX20-HPV18-E7, **slot 6** without, **slot 7** with heat-shock; **slot 8**: size markers; **slot 1** is exposed 10 times as long as **slots 2–7**. pORFEX10-HPV18-E7 gave a protein of identical position and intensity as in slot 3.

clones seems to reflect the occasional usage of a second ATG on a eukaryotic message (Kozak, 1978). No clones were obtained with the expression vector alone or, with a construct, where the CMV promoter points into the opposite direction from the Hm^r-gene.

Expression of the HPV 18 reading frame E7

Prokaryotic vectors that bear a strong and often regulated pro-

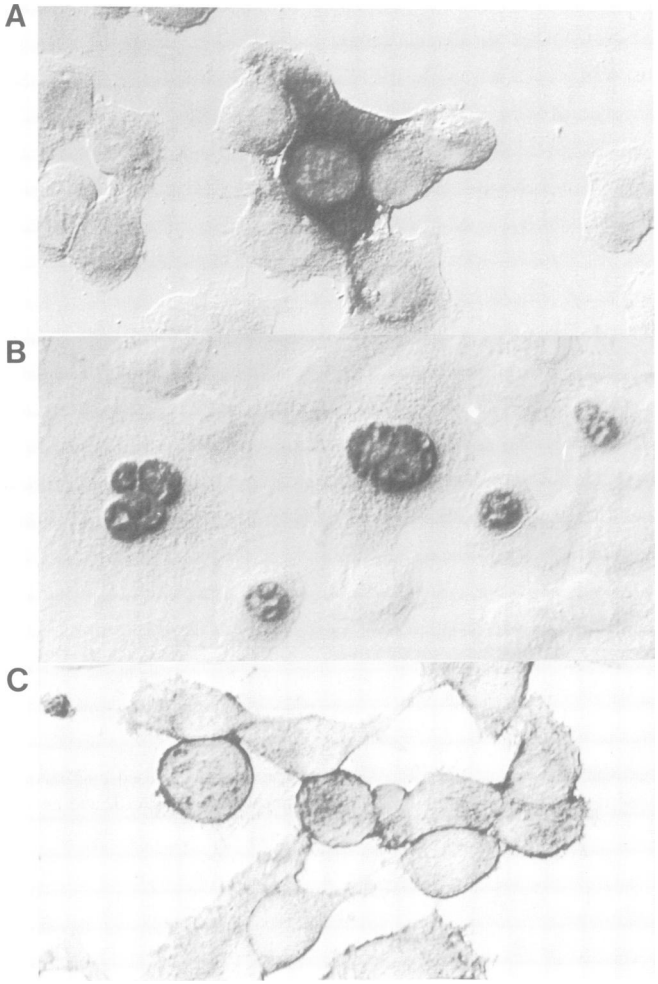

Fig. 5. Immunocytochemical visualization of HPV18-E7 protein in L cells. To contrast the cytoplasmically diffuse pattern obtained with HPV18-E7 **(A)**, L cells were independently transfected with a plasmid expressing SV40-TAg and stained with an anti TAg-antibody **(B)**. **(C)** cells stained with an anti $H2^k$ monoclonal antibody to exemplify plasma-membrane staining.

moter (e.g. Bernard *et al.*,1979) and determine the N-terminal part of a fusion protein (e.g. Remaut *et al.*, 1981) have been useful for the overexpression of proteins in *E. coli*. Vectors of this type were used for the expression of numerous ORFs of the HPV type 16 and 18 and rabbit antibodies against these fusion proteins were raised (Seedorf *et al.*, 1987). One of these antibodies was directed against a fusion protein containing at its N terminus 97 amino acids of the phage MS2 polymerase and its C terminus 94 amino acids of the HPV18 reading frame E7 (full size of E7: 105 triplets). The gene fusion was made with the help of an oligonucleotide linker with a *Bam*HI and an *Eco*RI site (Figure 3 and Seedorf *et al.*, 1987). In the *Bam*HI sequence, 5′-GGATCC-3′, the TCC triplet is in frame with the ORF of E7, a situation encountered at the *Bgl*II site of pORFEX10 and 20. In the *Eco*RI sequence, 5′-GAATTC-3′, the AAT triplet is in frame just as it is in pORFEX12, but not in pORFEX11 or 13. We inserted the E7-ORF as an 1100-bp *Bam*HI – *Hin*dIII or *Eco*RI – *Hin*dIII segment into these five vectors. L cells were transiently transfected, lysed and applied to a Western blot analysis. In the case of the construct pORFEX20-HPV18-E7, the cells were kept for 48 h at 37°C, and for 5 h at 42°C before lysis. Figure 4 shows that a 14-kd protein was detected by the rabbit anti MS2-HPV18-E7 antibody

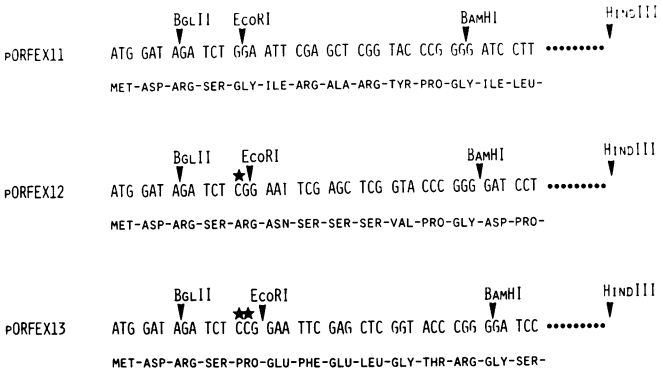

Fig. 6. N termini of fusion proteins expressed from pORFEX11, 12 and 13 as determined by the restriction site and the frameshift linker.

with the pORFEX10 and pORFEX12 constructs, but not with the 'frameshift-mutant-plasmids' pORFEX11 and 13. Also, no protein was detected in control cells and in cells transfected with a construct with opposite orientation of the CMV promoter. In the case of pORFEX20, the same band but with much weaker intensity was seen, and this only after heat-shock, not in the control cells kept at 37°C.

In HeLa cells which contain 10 – 50 endogenous copies of HPV18 (Schwarz *et al.*, 1985), the same antibody detected a smaller band of a mol. wt of 12.5 kd. This mol. wt is close to the size expected from the sequence which predicts an E7-protein with 105 amino acids in HeLa (Seedorf *et al.*, 1987; Seedorf, 1986), while pORFEX10-HPV18-E7 codes for one with 99 amino acids. Quantitative analysis by scintillation counting of E7 bands from a Western blot revealed that 100 times more E7 protein was expressed from pORFEX10-HPV18-E7 from the average L cell of the transfected plate than from the endogenous HPV18 copies of a HeLa cell (data not shown).

The HPV18-E8-protein is located in the cytoplasm of transfected L cells

To visualize the intracellular location of the E7 protein made from pORFEX10-HPV18-E7, immunocytochemistry of transiently transfected L cells was done. While ~50% of the cells were unstained like untransfected control cells, the other cells revealed a cytoplasmically diffuse staining pattern (Figure 5A). To exclude artefacts stemming from the transfection or immunocytochemical procedures, we transfected control cells with a plasmid containing the SV40 early region (pBSV3x; Banerji *et al.*, 1981) followed by reaction with the anti-SV40-LT monoclonal antibody PAb1619 (Ball *et al.*, 1984). Figure 5B shows the resulting nuclear staining pattern. Similarly, the monoclonal antibody H100-27.R55 (Lemke *et al.*, 1979) against the mouse $H\text{-}2^k$ protein revealed the location of this protein in this plasma membrane (Figure 5C). We conclude, that the majority, if not all, of the E7 protein expressed from pORFEX10-HPV18-E7 is located in the cytoplasm.

Discussion

We have designed and tested a plasmid vector system for the expression of genes and gene segments in transfected eukaryotic cells. In this publication in particular, we have aimed to document how a gene not containing its genuine promoter and nucleotide triplets coding for the N terminus of its protein product can be easily expressed from elements contributed by vec-

tor sequences. Beyond this, it was a particular strategy of our constructions to design this vector as a cassette-system where individual elements of a functionally understood and tested construct can be exchanged without disintegrating the whole genetic unit. To exemplify this, we exchanged a strong constitutive promoter against a regulatable heat shock promoter and compared their gene expression potential. Numerous restriction sites permit further modifications of this vector, e.g. to exchange the translation initiation region, or to obtain different N-terminal amino acids on fusion proteins such as peptides that would contain export signals (von Heijne, 1983). Figure 6 lists the N-terminal segments of proteins expressed after inserting genes into the singular *Eco*RI or *Bam*HI sites of the polylinker 3′ of the vector ATG. Further possible modifications include the elimination of the ATG for the expression of gene segments containing their genuine ATG, and of the insertion of translation termination signals into the *Hin*dIII site of these vectors. At this time, we do not present quantitative data on the expression of genes from pORFEX plasmids in stably transfected cells. But we are following a line of experiments where we maintain these plasmids on an episomal vector derived from bovine papillomavirus, pCGBPV9 (Matthias *et al.*, 1986) or Epstein – Barr virus, p205 (Yates *et al.*, 1985).

Our experiments are one approach within a project towards the analysis of the gene products of HPV. Numerous of the more than forty HPV types identified so far cause severe benign neoplastic lesions or are associated with malignant carcinomas (for reviews see Pfister, 1984; Gissmann, 1984). Many of these HPV types have very low levels of gene expression in natural biopsies and no cell culture systems for their amplification is broadly available (for recent progress see La Porta and Taichman, 1982 and Kreider *et al.*, 1985). These limitations reduce drastically our technical abilities to understand the relationship between the 10 or so ORFs common to all sequenced papillomavirus genomes (Chen *et al.*, 1982; Danos *et al.*, 1982; Schwarz *et al.*, 1983; Seedorf *et al.*, 1985) and the biological properties of their protein products. To approach this aim we have expressed the ORF E7 of HPV18 from our expression vectors. Western blot analysis with a rabbit anti-HPV18-E7 antibody of transiently transfected L cells reveals an ~100-fold higher expression of E7 protein from CMV promoter plasmids and 10-fold higher expression from the hsp promoter than found in HeLa cells, which contain ~10–50 endogenous HPV18 copies under the control of the HPV promoter (Schwarz *et al.*, 1985). Western blots with MS2-E7 protein expressed in *E. coli*, whose concentration was known, led to an estimate of the amount of E7 protein in HeLa cells close to 0.01% of total cell protein (Seedorf, 1985). Since only 50% of all L cell transfectants showed immune staining with anti-E7-antibody, and the total cell population had ~100-fold more E7 protein than HeLa cells, the average E7 positive transfectant should contain 2% of its total cell protein as E7 protein.

We do not know why this E7 protein shows a slightly higher molecular weight in transfected L cells than in HeLa cells. To exclude DNA cloning artefacts, we have sequenced by Maxam – Gilbert techniques a DNA segment between a *Bst*NI site within the CMV promoter segment and a *Hin*fI site in the E7 ORF. This experiment confirmed a continuous sequence spanning 20 3′ terminal nucleotides of the CMV promoter, 18 nucleotides from the pUC18 linker and the synthetic oligonucleotide containing the ATG and 36 nucleotides from the 5′ terminal part of E7 (data not shown) and proves the correct construction of a fusion gene as predicted by the DNA-ligation protocol outlined in this paper. The observation of at least two E7 bands in long exposures of Western blots of HeLa cell protein (not visible in Figure 4) suggests the possibility of modifications of the primary gene product, which may be different in mouse fibroblasts than in human epithelial cells. An unexpectedly low electrophoretic mobility was also observed for the recently identified protein E7 of HPV16 (Smotkin and Wettstein, 1986).

The anti-E7-antibody was not sufficiently sensitive to detect the E7 protein in HeLa cells fixed for immunostaining (K.Seedorf, unpublished). After overexpression of E7 from pORFEX vectors we found a diffuse immune staining in the cytoplasm of transfected L cells. It seems unlikely that this represents an unnatural pile-up of E7 protein at the site of translation, since cells with a low concentration of E7, e.g. after expression from the hsp promoter, showed a similar pattern of intracellular localization. We cannot exclude the possibility that the exchange of 11 N-terminal amino acids of E7 on the HPV18 genome by five amino acids coded by pORFEX vectors or the modifications that lead to the changed mol. wt also changed the biological properties of this protein. We suggest that pORFEX vectors should be a useful technique to approach the question for intracellular localization and biological function of further products from HPV ORFs and other genes of interest.

While research towards this publication was in progress a paper described a plasmid with a potential for similar purposes (Shyam *et al.*, 1986) without presenting data about its experimental application.

Materials and methods

DNA and DNA cloning procedures

All mentioned plasmids are either generally available or their source is mentioned in the text. All enzymes were from commercial suppliers and used according to the supplier's specifications. Oligonucleotides were synthesized with an Applied Biosystems 380 A DNA synthesizer and purified by protocols described in a manual from Applied Biosystems.

Culture, transfection and transfection analysis

Mouse LTk$^-$ fibroblasts were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal calf serum, 10 mM Hepes (pH 7.4) and 100 μg/ml penicillin/streptomycin. For transient DNA transfections 3×10^6 cells were plated out on a 10-cm dish, washed 24 h later twice with DMEM, then 10 μg of plasmid DNA were added in 2 ml DMEM containing 200 μg/ml DEAE – Dextran (0.5 Md, Sigma) and 0.1 mM Chloroquin. After 5 h at 37°C, the cells were washed twice with DMEM and then kept for 48 h in DMEM with the supplements given above. In case of the use of the hsp promoter, heat shock was for 5 h at 42°C at the end of this period. Analysis of transfected cells by CAT assay (Gorman *et al.*, 1982; 5 μg protein were reacted with 62.5 μCi [^{14}C]chloramphenicol) and Western-blotting (Burnette, 1981) followed published procedures. For immunocytochemistry cells were replated 24 h after DNA application on autoclaved glass slides. Twenty-four hours later the slides were washed in phosphate-buffered saline (PBS). The cells were fixed for 5 min in ice-cold methanol and for 2 min in acetone. Immunocytochemistry was performed using rabbit antisera against an MS2-HPV18-E7 fusion protein (Seedorf *et al.*, 1987; Seedorf, 1986) and a commercially available detection system (Amersham). The slides were first incubated for 1 h with the above anti-E7 antibody diluted 1:50–1:200, washed three times in PBS, incubated for 1 h with a biotinylated antirabbit antibody (diluted 1:50), washed again, incubated with Streptavidin-biotinylated peroxidase complexes (diluted 1:100) for 1 h.
3-Amino-9-ethylcarbazole (0.02%) in 50 mM acetate buffer pH 5.2 containing 0.93% H_2O_2 was used as substrate for the color reaction. All incubations were done at 37°C. Stable transfection and selection for resistance against hygromycin has been described by Bernard *et al.* (1985).

Acknowledgements

We wish to thank M.Boshart, G.Brandner, B.Gloss, B.Hovemann, G.Klock, R.Mestril, N.Koch, W.Röwekamp, W.Scheurlen and G.Schütz for materials and helpful discussions and M.Cole for excellent secretarial assistance.

References

Ball,R.K., Siegl,B., Quellhorst,S., Brandner,G. and Braun,D.G. (1984) *EMBO J.*, **3**, 1485–1491.

Banerji,J., Rusconi,S. and Schaffner,W. (1981) *Cell*, **27**, 299–308.
Bernard,H.U., Remaut,E., Hershfield,V., Das,H.K., Helinski,D.R., Yanofsky,C. and Franklin,N. (1979) *Gene*, **5**, 59–76.
Bernard,H.U., Krämmer,G. and Röwekamp,W.G. (1985) *Exp. Cell Res.*, **158**, 237–243.
Boshart,M., Weber,F., Jahn,G., Dorsch-Häsler,K., Fleckenstein,B. and Schaffner,W. (1985) *Cell*, **41**, 521–530.
Burnette,W.N. (1981) *Anal. Biochem.*, **112**, 195–203.
Chen,E.Y., Howley,P.M., Levinson,A.D. and Seeburg,P.H. (1982) *Nature*, **299**, 529–534.
Danos,O., Katinka,M. and Yamir,M. (1982) *EMBO J.*, **1**, 231–236.
Gissmann,L. (1984) *Cancer Surveys*, **3**, 161–181.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Hsiung,N., Fitts,R., Wilson,S., Milne,A. and Hamer,D. (1984) *J. Mol. Appl. Genet.*, **2**, 497–506.
Kozak,M. (1978) *Cell*, **15**, 1109–1123.
Kreider,J.W., Howett,M.K., Wolfe,S.A., Bartlett,G.L., Zaino,R.J., Sedlacek,T.V. and Mortel,R. (1985) *Nature*, **317**, 639–641.
La Porta,R.F. and Taichmann,L.B. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3393–3397.
Lemke,H., Hämmerling,G.J. and Hämmerling,U. (1979) *Immunol. Rev.*, **47**, 175–195.
Matthias,P., Boeger,U., Danesh,U., Schütz,G. and Bernard,H.U. (1986) *J. Mol. Biol.*, **187**, 557–568.
Mulligan,R. and Berg,P. (1980) *Science*, **209**, 1422–1427.
Nakabayashi,Y., Chattopadhyay,S.K. and Lowy,D.R. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5832–5836.
Okayama,H. and Berg,P. (1983) *Mol. Cell. Biol.*, **3**, 280–289.
Pfister,H. (1984) *Rev. Physiol. Biochem. Pharmacol.*, **99**, 111–181.
Remaut,E., Stanissens,P. and Fiers,W. (1981) *Gene*, **15**, 81–93.
Schwarz,E., Dürst,M., Demankowski,C., Lattermann,O., Zech,R., Wolfsperger,E., Suhai,S. and zur Hausen,H. (1983) *EMBO J.*, **2**, 2341–2348.
Schwarz,E., Freese,U.K., Gissmann,L., Mayer,W., Roggenbuck,B., Stremlan,A. and zur Hausen,H. (1985) *Nature*, **314**, 111–114.
Seedorf,K. (1986) Ph.D. Thesis, University of Heidelberg, FRG.
Seedorf *et al.* (1987) *EMBO J.*, **6**, 139–144.
Seedorf,K., Krämmer,G., Dürst,M., Suhai,S. and Röwekamp,W.G. (1985) *Virology*, **145**, 181–185.
Shyam,E., Reddy,P. and Rao,V.N. (1986) *Gene Anal. Techn.*, **3**, 41–44.
Smotkin,D. and Wettstein,F.O. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4680–4684.
Steller,H. (1985) Ph.D. Thesis, University of Heidelberg, FRG.
von Heijne,G. (1983) *Eur. J. Biochem.*, **133**, 17–21.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
Yates,J.L., Warren,N. and Sugden,B. (1985) *Nature*, **313**, 812–815.

Received on 1 September 1986; revised on 10 October 1986